Express Mail Label No.:EL251689298US

IN THE UNITED STATES PAPTENT AND TRADEMARK OFFICE

PROVISIONAL PATENT APPLICATION OF:

Kevin P. **LOWITZ**, Ruoping **CHEN**, Chen W. **LIAW**,

FOR:

NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED RECEPTOR:TDAG8

ARENA PHARMACEUTICALS, INC.
6166 Nancy Ridge Drive
San Diego, CA 92121

NON-ENDOGENOUS CONSTITUTIVELY ACTIVATED
HUMAN G PROTEIN-COUPLED RECEPTOR: TDAG8

TABLE OF CONTENTS

NON-ENDOGENOUS CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED RECEPTOR: TDAG8

The benefit of commonly owned U.S. Serial Number 09/170,496, filed via Express Mail on October 13, 1998, and Provisional Patent Application Serial Number 60/108,029, filed via Express Mail on November 12, 1998 is hereby claimed.

FIELD OF THE INVENTION

The invention disclosed in this patent document relates to transmembrane receptors, more particularly to endogenous, G protein-coupled receptor for which the endogenous ligand is unknown, and most particularly to the use of a non-endogenous, constitutively activated receptor, TDAG8, for the direct identification of candidate compounds via screening as agonists, partial agonists or inverse agonists to such receptors. Further disclosed are potential endogenous activators for TDAG8.

BACKGROUND OF THE INVENTION

A. G protein-coupled receptors

G protein-coupled receptors share a common structural motif. All these receptors have seven sequences of between 22 to 24 hydrophobic amino acids that form seven alpha helices, each of which spans the membrane. The transmembrane helices are joined by strands of amino acids having a larger loop between the fourth and fifth transmembrane helix on the extracellular side of the membrane. Another larger loop, composed primarily of hydrophilic amino acids, joins transmembrane helices five and six on the intracellular side of the membrane. The carboxy terminus of the receptor lies intracellularly with the amino terminus in the extracellular space. It is thought that the loop joining helices five and six, as

well as the carboxy terminus, interact with the G protein. Currently, Gq, Gs, Gi, and Go are G proteins that have been identified.

Under physiological conditions, G protein-coupled receptors exist in the cell membrane in equilibrium between two different states or conformations: an "inactive" state and an "active" state. A receptor in an inactive state is unable to link to the intracellular transduction pathway to produce a biological response. Changing the receptor conformation to the active state allows linkage to the transduction pathway and produces a biological response.

A receptor may be stabilized in an active state by an endogenous ligand or an exogenous agonist ligand. Recent discoveries such as, including but not exclusively limited to, modifications to the amino acid sequence of the receptor provide means other than ligands to stabilize the active state conformation. These means effectively stabilize the receptor in an active state by simulating the effect of a ligand binding to the receptor. Stabilization by such ligand-independent means is termed "constitutive receptor activation." A receptor for which the endogenous ligand is unknown or not identified is referred to as an "orphan receptor."

B. Traditional Compound Screening

Generally, the use of an orphan receptor for screening purposes to identify compounds that modulate a biological response associated with such receptor has not been possible. This is because the traditional "dogma" regarding screening of compounds mandates that the ligand for the receptor be known, whereby compounds that competitively bind with the receptor, *i.e.*, by interfering or blocking the binding of the natural ligand with the receptor, are selected. By definition, then, this approach has no applicability with

respect to orphan receptors. Thus, by adhering to this dogmatic approach to the discovery of therapeutics, the art, in essence, has taught and has been taught to forsake the use of orphan receptors unless and until the natural ligand for the receptor is discovered. The pursuit of an endogenous ligand for an orphan receptor can take several years and cost millions of dollars. Furthermore, and given that there are an estimated 2,000 G protein-coupled receptors in the human genome, the majority of which being orphan receptors, the traditional dogma castigates a creative approach to the discovery of therapeutics to these receptors.

The present invention relates to a human T-cell death-associated gene receptor (TDAG8). The deletion of self-reactive immature T-cells in the thymus is mediated by apoptosis upon T-cell receptor interaction. Apoptosis is characterized by a rapid collapse of the nucleus, extreme chromatin condensation, DNA fragmentation, and shrinkage of cells, and it is often dependent on the synthesis of new sets of RNA and protein. (*see,* Choi et al., 168 Cellular Immun. 78 (1996)). There is a strong correlation between apoptosis and TDAG8; *i.e.*, an increase in apoptosis results in an increase in the expression of TDAG8. *Id.* However, it is unknown whether an increase in TDAG8 expression causes T-cell mediated apoptosis, or if such expression is a result of such apoptosis.

The endogenous ligand for TDAG8 is unknown and is thus considered an orphan GPCR having an open reading frame of 1,011 bp encoding a 337 amino acid protein. (TDAG8 was cloned and sequenced in 1998. Kyaw, H. et al, 17 DNA Cell Biol. 493 (1998); *see* Figure 1 of Kyaw for nucleic and deduced amino acid sequences.). Human TDAG8 is reported to be homologous to murine TDAG8. Human TDAG8 is expressed in the liver and in lymphoid tissues, including peripheral blood leukocytes, spleen, lymph

nodes and thymus. TDAG8 is also reported to be localized to chromosome 14q31-32.1. *Id.*

SUMMARY OF THE INVENTION

Disclosed herein are methods for screening of candidate compounds against endogenous, constitutively activated G protein-coupled orphan receptor, most preferably against non-endogenous, constitutively activated G protein-coupled receptor, for the direct identification of candidate compounds as agonists, inverse agonists or partial agonists to TDAG8; and for biological detection of endogenous, constitutively activated G protein-coupled orphan receptor, TDAG8. For such screening purposes, it is preferred that a non-endogenous, constitutively activated GPCR, TDAG8, G protein-fusion protein be utilized.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a graphic representation of endogenous, constitutively activated TDAG8 ("TDAG8 WT") and non-endogenous, constitutively activated TDAG8 ("TDAG8 Mut") in a 293 cell-based cAMP assay.

Figure 2 is a graphic representation of ATP and ADP activation of CMV (control; expression vector) and endogenous, constitutively activated TDAG8 ("TDAG8 WT") in 293 cell-based cAMP assay.

Figure 3 is a graphic representation of ATP and ADP activation of CMV (control; expression vector) and endogenous, constitutively activated TDAG8 ("TDAG8 WT") grown in serum and serum-free media in 293 cell-based cAMP assay.

Figures 4A-4B is a representation of a dose response curve for endogenous, constitutively activated TDAG8 ("TDAG8 WT") in 293 cell-based cAMP assay. Figure 4A

shows ATP binding to "TDAG8 WT" at an EC50 value of 139.8uM, while Figure 4B shows ADP binding to "TDAG8 WT" at an EC50 value of 120.5uM.

Figures 5A-5B provides graphic results of comparative analysis of endogenous TDAG8 ("WT") versus non-endogenous, constitutively activated TDAG8 ("I225K") (control is designated "CMV") in 293 (5A) and 293T (5B) cells.

Figure 6 is a schematic representation of a portion of the preferred 8XCRE-LUC Reporter plasmid construct utilized herein.

Figure 7 is a reproduction of results of a tissue distribution of TDAG8 against various tissue-source mRNA's.

DETAILED DESCRIPTION

The scientific literature that has evolved around receptors has adopted a number of terms to refer to ligands having various effects on receptors. For clarity and consistency, the following definitions will be used throughout this patent document. To the extent that these definitions conflict with other definitions for these terms, the following definitions shall control:

AGONISTS shall mean materials (*e.g.*, ligands, candidate compounds) that activate the intracellular response when they bind to the receptor, or enhance GTP binding to membranes.

AMINO ACID ABBREVIATIONS used herein are set out in Table 1:

TABLE 1

ALANINE	ALA	A
ARGININE	ARG	R
ASPARAGINE	ASN	N

ASPARTIC ACID	ASP	D
CYSTEINE	CYS	C
GLUTAMIC ACID	GLU	E
GLUTAMINE	GLN	Q
GLYCINE	GLY	G
HISTIDINE	HIS	H
ISOLEUCINE	ILE	I
LEUCINE	LEU	L
LYSINE	LYS	K
METHIONINE	MET	M
PHENYLALANINE	PHE	F
PROLINE	PRO	P
SERINE	SER	S
THREONINE	THR	T
TRYPTOPHAN	TRP	W
TYROSINE	TYR	Y
VALINE	VAL	V

PARTIAL AGONISTS shall mean materials (*e.g.*, ligands, candidate compounds) which activate the intracellular response when they bind to the receptor to a lesser degree/extent than do agonists, or enhance GTP binding to membranes to a lesser degree/extent than do agonists

ANTAGONIST shall mean materials (*e.g.*, ligands, candidate compounds) that competitively bind to the receptor at the same site as the agonists but which do not activate the intracellular response initiated by the active form of the receptor, and can thereby inhibit the intracellular responses by agonists or partial agonists. ANTAGONISTS do not diminish the baseline intracellular response in the absence of an agonist or partial agonist.

CANDIDATE COMPOUND shall mean a molecule (for example, and not limitation, a chemical compound) which is amenable to a screening technique. Preferably, the phrase "candidate compound" does not include compounds which were publicly known to be compounds selected from the group consisting of inverse agonist, agonist or antagonist to a receptor, as previously determined by an indirect identification process ("indirectly identified compound"); more preferably, not including an indirectly identified compound which has previously been determined to have therapeutic efficacy in at least one mammal; and, most preferably, not including an indirectly identified compound which has previously been determined to have therapeutic utility in humans.

COMPOSITION means a material comprising at least one component; a "pharmaceutical composition" is an example of a composition.

COMPOUND EFFICACY shall mean a measurement of the ability of a compound to inhibit or stimulate receptor functionality, as opposed to receptor binding affinity. A most preferred means of detecting compound efficacy is via measurement of GTP (via [^{35}S]GTPγS) or cAMP, as further disclosed in the Example section of this patent document.

CONSTITUTIVELY ACTIVATED RECEPTOR shall mean a receptor subject to constitutive receptor activation. A constitutively activated receptor can be endogenous or non-endogenous.

CONSTITUTIVE RECEPTOR ACTIVATION shall mean stabilization of a receptor in the active state by means other than binding of the receptor with its endogenous ligand or a chemical equivalent thereof.

CONTACT or **CONTACTING** shall mean bringing at least two moieties together, whether in an in vitro system or an in vivo system.

DIRECTLY IDENTIFYING or **DIRECTLY IDENTIFIED**, in relationship to the phrase "candidate compound", shall mean the screening of a candidate compound against a constitutively activated receptor, preferably a constitutively activated orphan receptor, and most preferably against a constitutively activated G protein-coupled cell surface orphan receptor, and assessing the compound efficacy of such compound. This phrase is, under no circumstances, to be interpreted or understood to be encompassed by or to encompass the phrase "indirectly identifying" or "indirectly identified."

ENDOGENOUS shall mean a material that a mammal naturally produces. ENDOGENOUS in reference to, for example and not limitation, the term "receptor," shall mean that which is naturally produced by a mammal (for example, and not limitation, a human) or a virus. By contrast, the term **NON-ENDOGENOUS** in this context shall mean that which is not naturally produced by a mammal (for example, and not limitation, a human) or a virus. For example, and not limitation, a receptor which is not constitutively active in its endogenous form, but when manipulated becomes constitutively active, is most preferably referred to herein as a "non-endogenous, constitutively activated receptor." Both terms can be utilized to describe both "in vivo" and "in vitro" systems. For example, and not limitation, in a screening approach, the endogenous or non-endogenous receptor may be in reference to an in vitro screening system. As a further example and not limitation, where the genome of a mammal has been manipulated to include a non-endogenous constitutively activated receptor, screening of a candidate compound by means of an in vivo system is viable.

G PROTEIN COUPLED RECEPTOR FUSION PROTEIN and **GPCR FUSION PROTEIN**, in the context of the invention disclosed herein, each mean a non-

endogenous protein comprising an endogenous, constitutively activated orphan GPCR fused to at least one G protein, most preferably, the alpha (α) subunit of such G protein (this being the subunit that binds GTP), with the G protein preferably being of the same type as the G protein that naturally couples with endogenous orphan GPCR. For example, and not limitation, in an endogenous state, the G protein "Gsα" is the predominate G protein that couples with TDAG8 such that a GPCR Fusion Protein based upon TDAG8 would be a non-endogenous protein comprising TDAG8 fused to Gsα. The G protein can be fused directly to the c-terminus of the endogenous, constitutively active orphan GPCR or there may be spacers between the two.

INDIRECTLY IDENTIFYING or **INDIRECTLY IDENTIFIED** means the traditional approach to the drug discovery process involving identification of an endogenous ligand specific for an endogenous receptor, screening of candidate compounds against the receptor for determination of those which interfere and/or compete with the ligand-receptor interaction, and assessing the efficacy of the compound for affecting at least one second messenger pathway associated with the activated receptor.

INHIBIT or **INHIBITING,** in relationship to the term "response" shall mean that a response is decreased or prevented in the presence of a compound as opposed to in the absence of the compound.

INVERSE AGONISTS shall mean materials (*e.g.*, ligand, candidate compound) which bind to either the endogenous form of the receptor or to the constitutively activated form of the receptor, and which inhibit the baseline intracellular response initiated by the active form of the receptor below the normal base level of activity which is observed in the absence of agonists or partial agonists, or decrease GTP binding to membranes. Preferably,

the baseline intracellular response is inhibited in the presence of the inverse agonist by at least 30%, more preferably by at least 50%, and most preferably by at least 75%, as compared with the baseline response in the absence of the inverse agonist.

LIGAND shall mean an endogenous, naturally occurring molecule specific for an endogenous, naturally occurring receptor.

ORPHAN RECEPTOR shall mean an endogenous receptor for which the endogenous ligand specific for that receptor has not been identified or is not known.

PHARMACEUTICAL COMPOSITION shall mean a composition comprising at least one active ingredient, whereby the composition is amenable to investigation for a specified, efficacious outcome in a mammal (for example, and not limitation, a human). Those of ordinary skill in the art will understand and appreciate the techniques appropriate for determining whether an active ingredient has a desired efficacious outcome based upon the needs of the artisan.

NON-ORPHAN RECEPTOR shall mean an endogenous naturally occurring molecule specific for an endogenous naturally occurring ligand wherein the binding of a ligand to a receptor activates an intracellular signaling pathway.

STIMULATE or **STIMULATING**, in relationship to the term "response" shall mean that a response is increased in the presence of a compound as opposed to in the absence of the compound.

The order of the following sections is set forth for presentational efficiency and is not intended, nor should be construed, as a limitation on the disclosure or the claims to follow.

A. Introduction

The traditional study of receptors has always proceeded from the a priori assumption (historically based) that the endogenous ligand must first be identified before discovery could proceed to find antagonists and other molecules that could affect the receptor. Even in cases where an antagonist might have been known first, the search immediately extended to looking for the endogenous ligand. This mode of thinking has persisted in receptor research even after the discovery of constitutively activated receptors. What has not been heretofore recognized is that it is the active state of the receptor that is most useful for discovering agonists, partial agonists, and inverse agonists of the receptor. For those diseases which result from an overly active receptor, what is desired in a therapeutic drug is a compound which acts to diminish the active state of a receptor, not necessarily a drug which is an antagonist to the endogenous ligand. This is because a compound (drug) which reduces the activity of the active receptor state need not bind at the same site as the endogenous ligand. Thus, as taught by a method of this invention, any search for therapeutic compounds should start by screening compounds against the ligand-independent active state. The search, then, is for an inverse agonist to the active state receptor.

Screening candidate compounds against the endogenous, constitutively activated orphan receptor TDAG8, most preferably non-endogenous, constitutively activated orphan TDAG8, allows for the direct identification of candidate compounds which act at these orphan cell surface receptors, without requiring any prior knowledge or use of the receptor's endogenous ligand. By determining areas within the body where such receptors are expressed and/or over-expressed, it is possible to determine related disease/disorder states

which are associated with the expression and/or over-expression of these receptors; such an approach is disclosed in this patent document.

B. Disease/Disorder Identification and/or Selection

As will be set forth in greater detail below, most preferably inverse agonists or agonists to the non-endogenous, constitutively activated TDAG8 receptor can be identified by the methodologies of this invention. Such compound-types are ideal candidates as lead compounds in drug discovery programs for treating diseases related to this receptor. Because of the ability to directly identify inverse agonists and agonists to the TDAG8 receptor, thereby allowing for the development of pharmaceutical compositions, a search for diseases and disorders associated with the TDAG8 receptor is relevant. For example, scanning both diseased and normal tissue samples for the presence of the TDAG8 receptor now becomes more than an academic exercise or one which might be pursued along the path of identifying an endogenous ligand to TDAG8. Tissue scans can be conducted across a broad range of healthy and diseased tissues. Such tissue scans provide a preferred first step in associating a specific receptor with a disease and/or disorder.

Preferably, the DNA sequence of the TDAG8 receptor is used to make a probe for (a) dot-blot analysis against tissue-mRNA, and/or (b) RT-PCR identification of the expression of the receptor in tissue samples. The presence of a receptor in a tissue source, or a diseased tissue, or the presence of the receptor at elevated concentrations in diseased tissue compared to a normal tissue, can be preferably utilized to identify a correlation with a treatment regimen, including but not limited to, a disease associated with that disease. For example, TDAG8 is predominantly expressed in the lymphoid tissues, specifically the spleen, peripheral blood leukocytes and lymph nodes. Expression of TDAG8 has been

reported to increase during activation of-induced death of T-cell hybridomas stimulated by glucocorticoids or anti-T-cell receptor antibodies (*see,* Choi J.W. et al. 168 Cell. Immunol. 78 (1996)). This report suggests that TDAG8 may play a role in immature thymocyte deletion and peripheral T-cell development. Thus, an inverse agonist to TDAG8 is intended to prevent the death of T-cells upon activation, which is an important role in the human immune system. Receptors can equally well be localized to regions of organs by this technique. Based on the known functions of the specific tissues to which the receptor is localized, the putative functional role of the receptor can be deduced.

C. Screening of Candidate Compounds

1. *Generic GPCR screening assay techniques*

When a G protein receptor becomes constitutively active, it binds to a G protein (*e.g.,* Gq, Gs, Gi, Go) and stimulates the binding of GTP to the G protein. The G protein then acts as a GTPase and slowly hydrolyzes the GTP to GDP, whereby the receptor, under normal conditions, becomes deactivated. However, constitutively activated receptors continue to exchange GDP to GTP. A non-hydrolyzable analog of GTP, [^{35}S]GTPγS, can be used to monitor enhanced binding to membranes which express constitutively activated receptors. It is reported that [^{35}S]GTPγS can be used to monitor G protein coupling to membranes in the absence and presence of ligand. An example of this monitoring, among other examples well-known and available to those in the art, was reported by Traynor and Nahorski in 1995. The preferred use of this assay system is for initial screening of candidate compounds because the system is generically applicable to all G protein-coupled receptors regardless of the particular G protein that interacts with the intracellular domain of the receptor.

2. *Specific GPCR Screening Assay*

Once candidate compounds are identified using the "generic" G protein-coupled receptor assay (i.e. an assay to select compounds that are agonists, partial agonists, or inverse agonists), further screening to confirm that the compounds have interacted at the receptor site is preferred. For example, a compound identified by the "generic" assay may not bind to the receptor, but may instead merely "uncouple" the G protein from the intracellular domain. Adenylyl cyclase catalyzes the conversion of ATP to cAMP; thus, because these receptors are activated in their endogenous forms, increased levels of cAMP are associated therewith (on the other hand, endogenously activated receptors which couple the Gi protein are associated with decreased levels of cAMP). *See, generally,* "Indirect Mechanisms of Synaptic Transmission," Chpt. 8, From Neuron To Brain (3rd Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992).

A variety of approaches known in the art for measuring cAMP can be utilized; a most preferred approach relies upon the use of anti-cAMP antibodies. Another type of assay that can be utilized is a whole cell second messenger reporter system assay. Promoters on genes drive the expression of the proteins that a particular gene encodes. Cyclic AMP drives gene expression by promoting the binding of a cAMP-responsive DNA binding protein or transcription factor (CREB) which then binds to the promoter at specific sites called cAMP response elements and drives the expression of the gene. Reporter systems can be constructed which have a promoter containing multiple cAMP response elements before the reporter gene, e.g., β-galactosidase or luciferase. Thus, an activated Gs receptor such as TDAG8 causes the accumulation of cAMP which then activates the gene and expression of the reporter protein. The reporter protein such as β-galactosidase or

luciferase can then be detected using standard biochemical assays (Chen et al. 1995). A cAMP assay is particularly preferred.

The foregoing specific assay approach can, of course, be utilized to initially directly identify candidate compounds, rather than by using the generic assay approach. Such a selection is primarily a matter of choice of the artisan.

4. GPCR Fusion Protein

The use of an endogenous, constitutively activated orphan GPCR, most preferably a non-endogenous, constitutively activated orphan GPCR, for use in screening of candidate compounds for the direct identification of inverse agonists, agonists and partial agonists provides a unique challenge in that, by definition, the endogenous receptor is active even in the absence of an endogenous ligand bound thereto. Thus, in order to differentiate between, *e.g.,* the non-endogenous receptor in the presence of a candidate compound and the non-endogenous receptor in the absence of that compound, with an aim of such a differentiation to allow for an understanding as to whether such compound may be an inverse agonist, agonist, partial agonist or have no affect on such a receptor, it is preferred that an approach be utilized that can enhance such differentiation. A preferred approach is the use of a GPCR Fusion Protein.

Generally, once it is determined that a non-endogenous orphan GPCR, such as TDAG8, has been constitutively activated, using the assay techniques set forth above (as well as others), it is possible to determine the predominant G protein that couples with the endogenous GPCR. Coupling of the G protein to the GPCR provides a signaling pathway that can be assessed. Because it is most preferred that screening take place by use of a mammalian expression system, such a system will be expected to have endogenous G protein therein. Thus, by definition, in such a system, the non-endogenous, constitutively

active orphan GPCR will continuously signal. In this regard, it is preferred that this signal be enhanced such that in the presence of, *e.g.*, an inverse agonist to the receptor, it is more likely that on will be able to more readily differentiate, particularly in the context of screening, between the receptor when it is contacted with the inverse agonist.

The GPCR Fusion Protein is intended to enhance the efficacy of G protein coupling with the non-endogenous GPCR. The GPCR Fusion Protein is preferred for screening with a non-endogenous, constitutively activated GPCR because such an approach increases the signal that is most preferably utilized in such screening techniques. This is important in facilitating a significant "signal to noise" ratio; such a significant ratio is import preferred for the screening of candidate compounds as disclosed herein.

The construction of a construct useful for expression of a GPCR Fusion Protein is within the purview of those having ordinary skill in the art. Commercially available expression vectors and systems offer a variety of approaches that can fit the particular needs of an investigator. The criteria of importance for such a GPCR Fusion Protein construct is that the endogenous GPCR sequence and the G protein sequence both be in-frame (preferably, the sequence for the endogenous GPCR is upstream of the G protein sequence) and that the "stop" codon of the GPCR must be deleted or replaced such that upon expression of the GPCR, the G protein can also be expressed. The GPCR can be linked directly to the G protein, or there can be spacer residues between the two (preferably, no more than about 12, although this number can be readily ascertained by one of ordinary skill in the art). We have a preference (based upon convenience) of use of a spacer in that some restriction sites that are not used will, effectively, upon expression, become a spacer. Most preferably, the G protein that couples to the non-

endogenous GPCR will have been identified prior to the creation of the GPCR Fusion Protein construct. Because there are only a few G proteins that have been identified, it is preferred that a construct comprising the sequence of the G protein (*i.e.*, a universal G protein construct) be available for insertion of an endogenous GPCR sequence therein; this provides for efficiency in the context of large-scale screening of a variety of different endogenous GPCRs having different sequences.

D. cAMP Detection Assay

TDAG8 has been discovered to contain a conserved motif commonly found in purinergic receptors (*e.g.*, human P2Y). Purinoceptors contain conserved residues with positively charged amino acids (*e.g.*, His and Arg) and are preferentially activated by adenosine nucleotides (*e.g.*, ATP and ADP). Communi et al., 272 Jo. of Biol. Chem. 31969 (1997). Thus, the binding of adenosine nucleotides to purinoceptors can be coupled to the stimulation of adenylyl cyclase. Although TDAG8 is not characterized as a purinoceptor, the common motif, located before the "DRY" region of a GPCR, led us to determine whether ATP and/or ADP are potential endogenous activators of TDAG8.

In the case of TDAG8, it has been determined that this receptor couples the G protein Gs. Gs is known to activate the enzyme adenylyl cyclase, which is necessary for catalyzing the conversion of ATP to cAMP. Although no known endogenous ligand has been identified for TDAG8, such that TDAG8 is considered an orphan GPCR, Figure 2 evidences that ATP and ADP bind to TDAG8, resulting in an increase in cAMP. From this data, both of the adenosine nucleotides act as endogenous activators to TDAG8, and as endogenous activators, they increase the level of cAMP about 59% and about 55%, respectively.

Although a variety of cells are available to the art for the expression of proteins, it is most preferred that mammalian cells be utilized. The primary reason for this is predicated upon practicalities, *i.e.*, utilization of, *e.g.*, yeast cells for the expression of a GPCR, while possible, introduces into the protocol a non-mammalian cell which may not (indeed, in the case of yeast, does not) include the receptor-coupling, genetic-mechanism and secretary pathways that have evolved for mammalian systems – thus, results obtained in non-mammalian cells, while of potential use, are not as preferred as that obtained from mammalian cells. Of the mammalian cells, COS-7, 293 and 293T cells are preferred, with 293 cells being particularly preferred, although the specific mammalian cell utilized can be predicated upon the particular needs of the artisan.

E. Medicinal Chemistry

Generally, but not always, direct identification of candidate compounds is preferably conducted in conjunction with compounds generated via combinatorial chemistry techniques, whereby thousands of compounds are randomly prepared for such analysis. Generally, the results of such screening will be compounds having unique core structures; thereafter, these compounds are preferably subjected to additional chemical modification around a preferred core structure(s) to further enhance the medicinal properties thereof. In this way, inverse agonists, agonists and/or partial agonists that are directly identified can be beneficially improved prior to development of pharmaceutical compositions comprising such compounds. Generally, it is preferred that the binding affinity of a directly identified compound selected for further refinement into a pharmaceutical composition have a binding affinity for the receptor of less than 100nM, although this is generally a preference selection based upon the particular needs of the

artisan. Such techniques are known to those in the art and will not be addressed in detail in this patent document.

E. Pharmaceutical compositions

Candidate compounds selected for further development can be formulated into pharmaceutical compositions using techniques well known to those in the art. Suitable pharmaceutically-acceptable carriers are available to those in the art; for example, see Remington's Pharmaceutical Sciences, 16th Edition, 1980, Mack Publishing Co., (Oslo et al., eds.).

F. Other Utility

Although a preferred use of the non-endogenous versions of human TDAG8 is for the direct identification of candidate compounds as inverse agonists, agonists or partial agonists (preferably for use as pharmaceutical agents), this version of human TDAG8 can also be utilized in research settings. For example, *in vitro* and *in vivo* systems incorporating TDAG8 can be utilized to further elucidate and understand the role(s) TDAG8 plays in the human condition, both normal and diseased, as well as understanding the role of constitutive activation as it applies to understanding the signaling cascade. A value in non-endogenous human TDAG8 is that its utility as a research tool is enhanced in that, because of its unique features, non-endogenous TDAG8 can be used to understand the role of TDAG8 in the human body before the endogenous ligand therefor is identified. Other uses of the disclosed receptors will become apparent to those in the art based upon, *inter alia*, a review of this patent document.

EXAMPLES

The following examples are presented for purposes of elucidation, and not limitation, of the present invention. While specific nucleic acid and amino acid sequences are disclosed herein, those of ordinary skill in the art are credited with the ability to make minor modifications to these sequences while achieving the same or substantially similar results reported below.

Example 1
PREPARATION OF ENDOGENOUS, CONSTITUTIVELY ACTIVATED TDAG8

PCR was performed using genomic DNA as template and rTth polymerase (Perkin Elmer) with the buffer system provided by the manufacturer, 0.25 μM of each primer, and 0.2 mM of each 4 nucleotides. The cycle condition was 30 cycles of 94°C for 1 min, 56°C for 1min and 72 °C for 1 min and 20 sec. The 5' PCR primer contained a HindIII site with the following sequence:

5'-TGCAAGCTTAAAAAGGAAAAAATGAACAGC-3' (SEQ.ID.NO.:1)

and the 3' primer contained a BamHI site with the following sequence:

5'-TAAGGATCCCTTCCCTTCAAAACATCCTTG -3' (SEQ.ID.NO.:2).

The resulting 1.1 kb PCR fragment was digested with HindIII and BamHI and cloned into HindIII-BamHI site of pCMV expression vector. Three resulting clones sequenced contained three potential polymorphisms involving changes of amino acid 43 from Pro to Ala, amino acid 97 from Lys to Asn and amino acid 130 from Ile to Phe. Nucleic acid (SEQ.ID.NO.:3) and amino acid (SEQ.ID.NO.:4) sequences for human TDAG8 were thereafter determined.

Example 2
PREPARATION OF NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED TDAG8

Preparation of the non-endogenous, constitutively activated human TDAG8 receptor was accomplished by creating a I225K mutation (see, SEQ.ID.NO.:5 for nucleic acid sequence, SEQ.ID.NO.:6 for amino acid sequence). Mutagenesis was performed using Transformer Site-Directed™ Mutagenesis Kit (Clontech) according to manufacturer's instructions. The two mutagenesis primers were utilized, a lysine mutagenesis oligonucleotide (SEQ.ID.NO.:7) and a selection marker oligonucleotide (SEQ.ID.NO.:8), which had the following sequences:

5'- GGAAAAGAAGAGAATCAAAAAACTACTTGTCAGCATC -3' (SEQ.ID.NO.: 7)

5'- CTCCTTCGGTCCTCCTATCGTTGTCAGAAGT -3' (SEQ.ID.NO.: 8),

respectively.

Example 3
GPCR Fusion Protein Preparation

The design of the constitutively activated GPCR-G protein fusion construct was accomplished as follows: both the 5' and 3' ends of the rat G protein Gsα (long form; Itoh, H. et al., 83 PNAS 3776 (1986)) were engineered to include a HindIII (5'-AAGCTT-3') sequence thereon. Following confirmation of the correct sequence (including the flanking HindIII sequences), the entire sequence was shuttled into pcDNA3.1(-) (Invitrogen, cat. no. V795-20) by subcloning using the HindIII restriction site of that vector. The correct orientation for the Gsα sequence was determined after subcloning into pcDNA3.1(-). The modified pcDNA3.1(-) containing the rat Gsα gene at HindIII sequence was then verified; this vector was now available as a "universal" Gsα protein vector. The pcDNA3.1(-) vector contains a variety of well-known restriction sites

upstream of the HindIII site, thus beneficially providing the ability to insert, upstream of the Gs protein, the coding sequence of an endogenous, constitutively active GPCR. This same approach can be utilized to create other "universal" G protein vectors, and, of course, other commercially available or proprietary vectors known to the artisan can be utilized – the important criteria is that the sequence for the GPCR be upstream and in-frame with that of the G protein.

TDAG8-Gsα Fusion Protein construct was then made as follows: primers were designed for both endogenous, constitutively activated and non-endogenous, constitutively activated TDAG8 were as follows:

5'-gatcTCTAGAATGAACAGCACATGTATTGAAG-3' (SEQ.ID.NO.9; sense)

5'-ctagGGTACCCGCTCAAGGACCTCTAATTCCATAG-3' (SEQ.ID.NO.10; antisense).

Nucleotides in lower caps are included as spacers in the restriction sites between the G protein and TDAG8. The sense and anti-sense primers included the restriction sites for XbaI and KpnI, respectively.

PCR was then utilized to secure the respective receptor sequences for fusion within the Gsα universal vector disclosed above, using the following protocol for each: 100ng cDNA for TDAG8 was added to separate tubes containing 2ul of each primer (sense and anti-sense), 3uL of 10mM dNTPs, 10uL of 10XTaqPlus™ Precision buffer, 1uL of TaqPlus™ Precision polymerase (Stratagene: #600211), and 80uL of water. Reaction temperatures and cycle times for TDAG8 were as follows: the initial denaturing step was done it 94°C for five minutes, and a cycle of 94°C for 30 seconds; 55°C for 30 seconds; 72°C for two minutes. A final extension time was done at 72°C for ten minutes. PCR

product for was run on a 1% agarose gel and then purified (data not shown). The purified product was digested with XbaI and KpnI (New England Biolabs) and the desired inserts will be isolated, purified and ligated into the Gs universal vector at the respective restriction site. The positive clones was isolated following transformation and determined by restriction enzyme digest; expression using 293 cells was accomplished following the protocol set forth *infra*. Each positive clone for TDAG8:Gs – Fusion Protein was sequenced and made available for the direct identification of candidate compounds.

GPCR Fusion Proteins comprising the endogenous, constitutively activated TDAG8 and the non-endogenous, constitutively activated TDAG8 (I225K) will be analyzed as above and verified for constitutive activation (data not shown).

Example 4
REPORTER -BASED ASSAY: CRE-LUC REPORTER ASSAY

293 and 293T cells were plated-out on 96 well plates at a density of 2 x 10^4 cells per well and were transfected using Lipofectamine Reagent (BRL) the following day according to manufacturer instructions. A DNA/lipid mixture was prepared for each 6-well transfection as follows: 260ng of plasmid DNA in 100μl of DMEM were gently mixed with 2μl of lipid in 100μl of DMEM (the 260ng of plasmid DNA consisted of 200ng of a 8xCRE-Luc reporter plasmid (*see* below and Figure 2 for a representation of a portion of the plasmid), 50ng of pCMV comprising endogenous receptor or non-endogenous receptor or pCMV alone, and 10ng of a GPRS expression plasmid (GPRS in pcDNA3 (Invitrogen)). The 8XCRE-Luc reporter plasmid (Figure 5) was prepared as follows: vector SRIF-β-gal was obtained by cloning the rat somatostatin promoter (-71/+51) at BglV-HindIII site in the pβgal-Basic Vector (Clontech). Eight (8) copies of cAMP response element were obtained by PCR from an adenovirus template

AdpCF126CCRE8 (*see* 7 Human Gene Therapy 1883 (1996)) and cloned into the SRIF-β-gal vector at the Kpn-BglV site, resulting in the 8xCRE-β-gal reporter vector. The 8xCRE-Luc reporter plasmid was generated by replacing the beta-galactosidase gene in the 8xCRE-β-gal reporter vector with the luciferase gene obtained from the pGL3-basic vector (Promega) at the HindIII-BamHI site. Following 30 min. incubation at room temperature, the DNA/lipid mixture was diluted with 400 µl of DMEM and 100µl of the diluted mixture was added to each well. 100 µl of DMEM with 10% FCS were added to each well after a 4hr incubation in a cell culture incubator. The following day the transfected cells were changed with 200 µl/well of DMEM with 10% FCS. Eight (8) hours later, the wells were changed to 100 µl /well of DMEM without phenol red, after one wash with PBS. Luciferase activity were measured the next day using the LucLite™ reporter gene assay kit (Packard) following manufacturer instructions and read on a 1450 MicroBeta™ scintillation and luminescence counter (Wallac). Results are summarized in Figures 5A and 5B.

Example 4
CELL-BASED DETECTION ASSAY

293 cells were plated-out on 150mm plates at a density of 1.3 x 10^7 cells per plate, and were transfected using 12ug of the respective DNA and 60ul of Lipofectamine Reagent (BRL) per plate. The transfected cells were grown in media containing serum for an assay performed 24 hours post-transfection. For detection assay performed 48 hours post-transfection (assay comparing serum and serum-free media; see Figure 3), the initial media was changed to either serum or serum-free media. The serum-free media was comprised solely of Dulbecco's Modified Eagle's (DME) High Glucose Medium (Irvine

Scientific #9024). In addition to the above DME Medium, the media with serum contained the following: 10% Fetal Bovine Serum (Hyclone #SH30071.03), 1% of 100mM Sodium Pyruvate (Irvine Scientific #9334), 1% of 20mM L-Glutamine (Irvine Scientific #9317), and 1% of Penicillin-Streptomycin solution (Irvine Scientific #9366).

A 96-well Adenylyl Cyclase Activation Flashplate was used (NEN: #SMP004A). First, 50ul of the standards for the assay were added to the plate, in duplicate, ranging from concentrations of 50pmol to zero pmol cAMP per well. The standard cAMP (NEN: #SMP004A) was reconstituted in water, and serial dilutions were made using 1xPBS (Irvine Scientific: #9240). Next, 50ul of the stimulation buffer (NEN: #SMP004A) was added to all wells. In the case of using compounds to measure activation or inactivation of cAMP, 10ul of each compound, diluted in water, was added to its respective well, in triplicate. Various final concentrations used range from 1uM up to 1mM. Adenosine 5'-triphosphate, ATP, (Research Biochemicals International: #A-141) and Adenosine 5'-diphosphate, ADP, (Sigma: #A2754) were used in the assay. Next, the 293 cells transfected with the respective cDNA (CMV or TDAG8) were harvested 24 (assay detection in serum media) or 48 hours post-transfection (assay detection comparing serum and serum-free media). The media was aspirated and the cells washed once with 1xPBS. Then 5ml of 1xPBS was added to the cells along with 3ml of cell dissociation buffer (Sigma: #C-1544). The detached cells were transferred to a centrifuge tube and centrifuged at room temperature for five minutes. The supernatant was removed and the cell pellet was resuspended in an appropriate amount of 1xPBS to obtain a final concentration of $2x10^6$ cells per milliliter. To the wells containing the compound, 50ul of the cells in 1xPBS ($1x10^5$ cells/well) were added. The plate was incubated on a shaker for 15 minutes at room temperature. The detection buffer containing

the tracer cAMP was prepared. In 11ml of detection buffer (NEN: #SMP004A), 50ul (equal to 1uCi) of [^{125}I]cAMP (NEN: #SMP004A) was added. Following incubation, 50ul of this detection buffer containing tracer cAMP was added to each well. The plate was placed on a shaker and incubated at room temperature for two hours. Finally, the solution from the wells of the plate were aspirated and the flashplate was counted using the Wallac MicroBeta scintillation counter.

In Figure 2 ATP and ADP bind to endogenous TDAG8 resulting in an increase of cAMP of about 59% and about 55% respectively. Figure 3 evidences ATP and ADP binding to endogenous TDAG8 where endogenous TDAG8 was transfected and grown in serum and serum-free medium. ATP binding to endogenous TDAG8 grown in serum media evidences an increase in cAMP of about 65%, compared to the endogenous TDAG8 with no compounds; in serum-free media there was an increase of about 68%. ADP binding to endogenous TDAG8 in serum evidences about a 61% increase, while in serum-free ADP binding evidences an increase of about 62% increase. In Figures 4A and 4B, ATP and ADP bind to endogenous TDAG8 with an EC50 value of 139.8uM and 120.5uM, respectively.

Although the results presented in Figure 3 indicate substantially the same results when serum and serum-free media were compared, our choice is to use a serum based media, although a serum-free media can also be utilized.

Example 5
TISSUE DISTRIBUTION OF TDAG8

Before using a multiple tissue cDNA ("MTC") panel, two primers were designed from the TDAG8 open reading frame sequence. The oligonucleotides utilized were as follows:

5'-GCACTCATGGTCAGCCTGTCCATC-3' (SEQ.ID.NO.:11; sense),

5'-GTACAGAATTGGATCAGCAACAC-3' (SEQ.ID.NO.:12; antisense).

Once the two primers were made and purified for PCR use, the reaction mixes were made. Each tube contained the following master mix of reagents: 36ul water, 1ul 10mM dNTP mix, 1ul Taq Plus Precision DNA Polymerase (Stratagene: #600211), and 5ul 10x Buffer for Taq Plus Precision Polymerase (Stratagene: #600211). A positive control tube containing the above solutions also contained 2ul of the G3PDH positive control primers (Clontech: #K14261-1) and 5ul of the control cDNA (Clontech: #K1426-1). A negative control tube was similar to the positive control; however, the control cDNA was replaced with 5ul water. To determine gene distribution, the MTC Panels used included the Human Panel I (Clontech: #K1420-1), the Human Panel II (Clontech: #K1421-1), and the Human Immune System Panel (Clontech: #K1426-1). Each MTC Panel contained several tubes of cDNA from various human tissues. Using tubes containing the above master mix of reagents, 2ul of the G3PDH positive control primers, 5ul of the individual MTC Panel cDNA, and 1ul of each primer designed above for the TDAG8 gene were all added to complete the reaction mixture. All of the tubes were then placed into a programmable thermal cycler (Perkin Elmer). The reactions were added at 94C for 30 seconds. A cycle of 94C for 30 seconds, 55C for 30 seconds, and 72C for two minutes was repeated 30 times. A final extension time of five minutes at 72C was run, and the cooling temperature of 4C was the final step. The reactions were added to a one percent agarose gel (FMC Bioproducts: #50004) and examined under ultraviolet light. For the positive control, an expected band of approximately 1Kb should be seen. For the negative control, no band should appear. Finally, for all of the tubes containing various

human tissue cDNA, a band should be seen at 1Kb (for the control primers), and if expressed in that particular tissue, a band (varying in intensity) should be seen at 450bp (for the TDAG8 primers). *See* Figure 7.

References cited throughout this patent document, unless otherwise indicated, are incorporated herein by reference. Modifications and extension of the disclosed inventions that are within the purview of the skilled artisan are encompassed within the above disclosure and the claims that follow.

Although a variety of expression vectors are available to those in the art, for purposes of utilization for the endogenous and non-endogenous human TDAG8, as well as the GPCR Fusion Protein comprising endogenous and non-endogenous TDAG8, it is most preferred that the vector utilized be pCMV. This vector was deposited with the American Type Culture Collection (ATCC) on October 13, 1998 (10801 University Blvd., Manassas, VA 20110-2209 USA) under the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure. The DNA was tested by the ATCC and determined to be. The ATCC has assigned the following deposit number to pCMV: ATCC #203351.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Lowitz, Kevin P.
Chen, Ruoping
Liaw, Chen W.

(ii) TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Human G Protein-Coupled Receptor: TDAG8

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Arena Pharmaceuticals, Inc.
(B) STREET: 6166 Nancy Ridge Drive
(C) CITY: San Diego
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92121

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Burgoon, Richard P.
(B) REGISTRATION NUMBER: 34,787

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (858)453-7200
(B) TELEFAX: (858)453-7210

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGCAAGCTTA AAAAGGAAAA AATGAACAGC 30

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TAAGGATCCC TTCCCTTCAA AACATCCTTG 30

(4) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT 60

TACATCTTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTG TGTGTCTTTC 120

CTGCAACCCA AGAAGGAAAG TGAACTAGGA ATTTACCTCT TCAGTTTGTC ACTATCAGAT 180

TTACTCTATG CATTAACTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACTGG 240

ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTTCTCA TGTACATGAA GTTTTACAGC 300

AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCTTTG 360

AAGTTTTTTT TCCTAAGGAC AAGAAGAATT GCACTCATGG TCAGCCTGTC CATCTGGATA 420

TTGGAAACCA TCTTCAATGC TGTCATGTTG TGGGAAGATG AAACAGTTGT TGAATATTGC 480

GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA 540

ATCAACCTCA ACTTGTTCAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCATCCTG 600

ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGA AAACAAGGAA 660

AAGAAGAGAA TCATAAAACT ACTTGTCAGC ATCACAGTTA CTTTTGTCTT ATGCTTTACT 720

CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTTAGAGC ATGCTGTGAA CTTCGAAGAC 780

CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT 840

TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTTTTGTTA CCGAAACAGG AAGATATGAT 900

ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAGAAAA 960

CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCTTGA GTAG 1014

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
1 5 10 15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
20 25 30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Pro Lys Lys Glu Ser Glu
35 40 45

Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
50 55 60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65 70 75 80

Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
85 90 95

Lys Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
115 120 125

Arg Ile Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
130 135 140

Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
145 150 155 160

Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
165 170 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
180 185 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
195 200 205

Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
210 215 220

Ile Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
225 230 235 240

Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
245 250 255

Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
260 265 270

Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
275 280 285

Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
290 295 300

Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305 310 315 320

Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
325 330 335

Glu

(6) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1014 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT 60

TACATCTTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTG TGTGTCTTTC 120

CTGCAAGCAA AGAAGGAAAG TGAACTAGGA ATTTACCTCT TCAGTTTGTC ACTATCAGAT 180

TTACTCTATG CATTAACTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACTGG 240

ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTTCTCA TGTACATGAA TTTTTACAGC 300

AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCTTTG 360

AAGTTTTTTT TCCTAAGGAC AAGAAGATTT GCACTCATGG TCAGCCTGTC CATCTGGATA 420

TTGGAAACCA TCTTCAATGC TGTCATGTTG TGGGAAGATG AAACAGTTGT TGAATATTGC 480

GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA 540

ATCAACCTCA ACTTGTTCAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCATCCTG 600

ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGA AAACAAGGAA 660

AAGAAGAGAA TCAAAAAACT ACTTGTCAGC ATCACAGTTA CTTTTGTCTT ATGCTTTACT 720

CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTTAGAGC ATGCTGTGAA CTTCGAAGAC 780

CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT 840

TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTTTTGTTA CCGAAACAGG AAGATATGAT 900

ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAGAAAA 960

CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCTTGA GTAG 1014

(7) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu
1 5 10 15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn
20 25 30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Ala Lys Lys Glu Ser Glu
35 40 45

Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala
50 55 60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp
65 70 75 80

Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met
85 90 95

Asn Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg
100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Phe Leu Arg Thr Arg
115 120 125

Arg Phe Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile
130 135 140

Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys
145 150 155 160

Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu
165 170 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr
180 185 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln
195 200 205

Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile
210 215 220

Lys Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr
225 230 235 240

Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val
245 250 255

Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr
260 265 270

Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile
275 280 285

Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile
290 295 300

Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys
305 310 315 320

Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu
325 330 335

Glu

(8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAAAAGAAG AGAATCAAAA AACTACTTGT CAGCATC 37

(9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T 31

(10) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATCTCTAGA ATGAACAGCA CATGTATTGA AG 32

(11) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTAGGGTACC CGCTCAAGGA CCTCTAATTC CATAG 35

(12) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCACTCATGG TCAGCCTGTC CATC 24

(13) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTACAGAATT GGATCAGCAA CAC 23

ABSTRACT

The invention disclosed in this patent document relates to transmembrane receptors, more particularly to G protein-coupled receptors for which the endogenous ligand is unknown ("orphan GPCR receptors"), and most particularly to a mutated (non-endogenous) version of human TDAG8, with such mutated version being constitutively active.

293 Cell-Based cAMP Assay

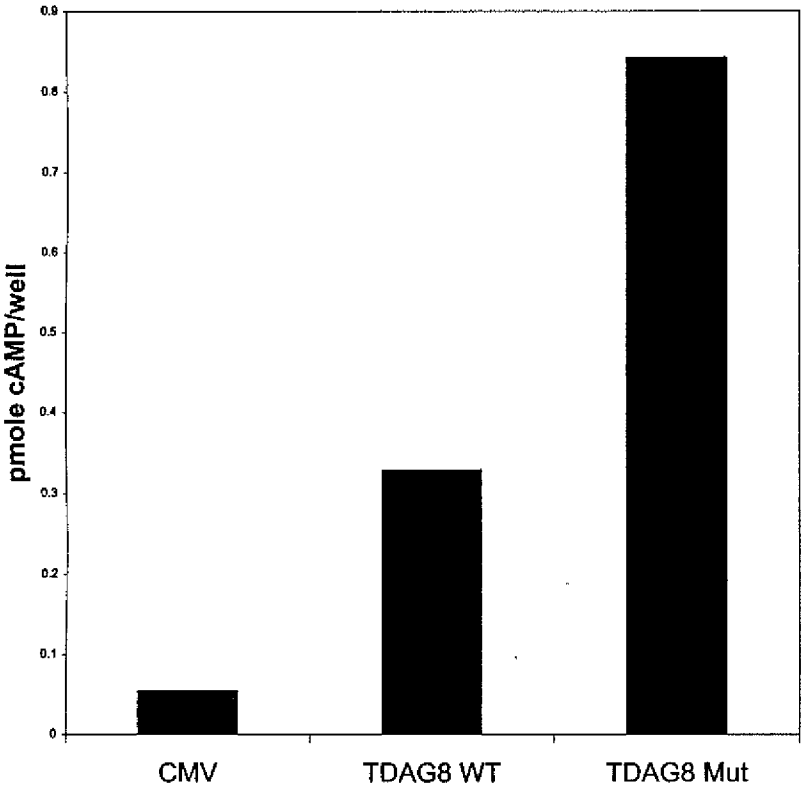


FIGURE 1

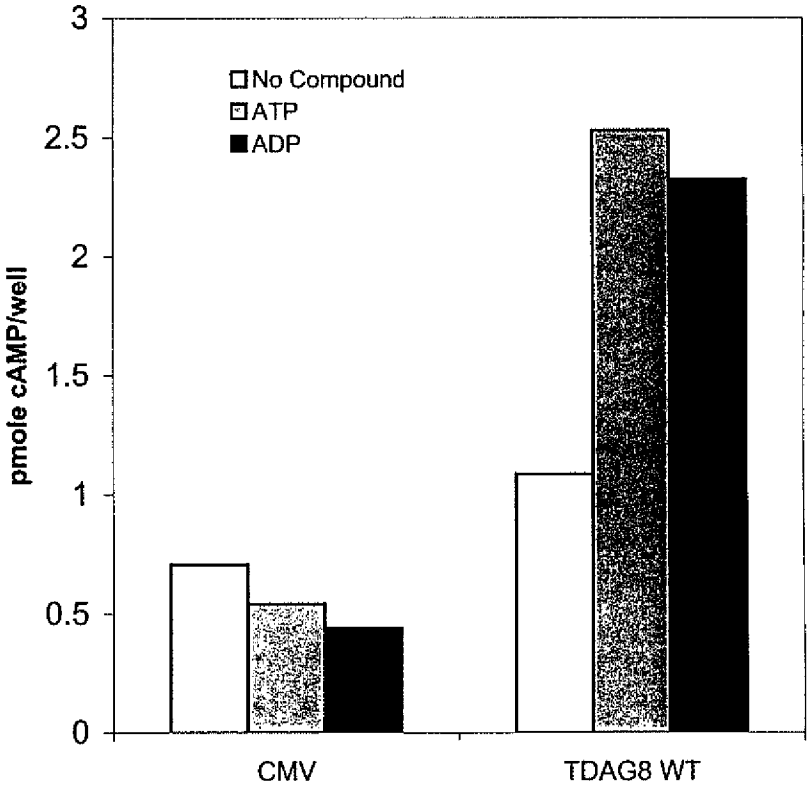


FIGURE 2

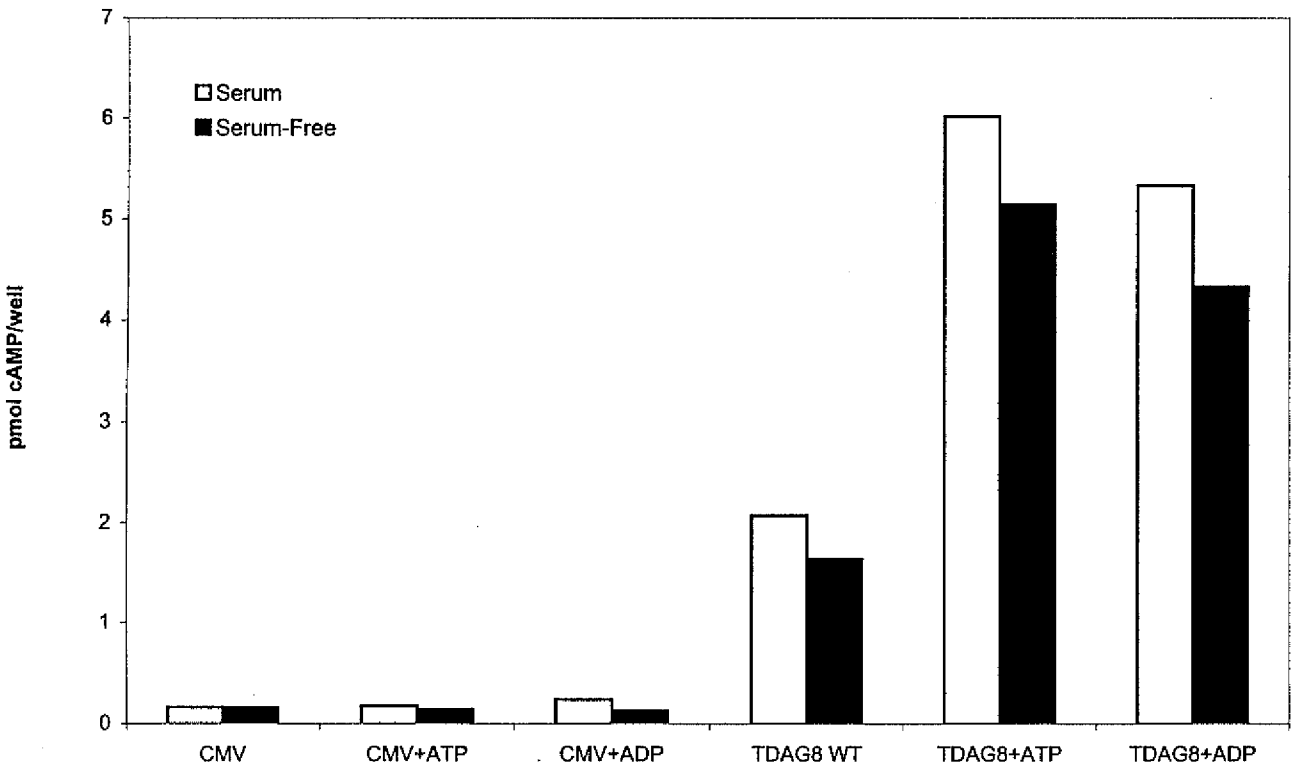

293 Cell-Based cAMP Assay
pmol cAMP/well
Serum
Serum-Free
0
1
2
3
4
5
6
7
CMV
CMV+ATP
CMV+ADP
TDAG8 WT
TDAG8+ATP
TDAG8+ADP

FIGURE 3

293 Cell-based cAMP Assay

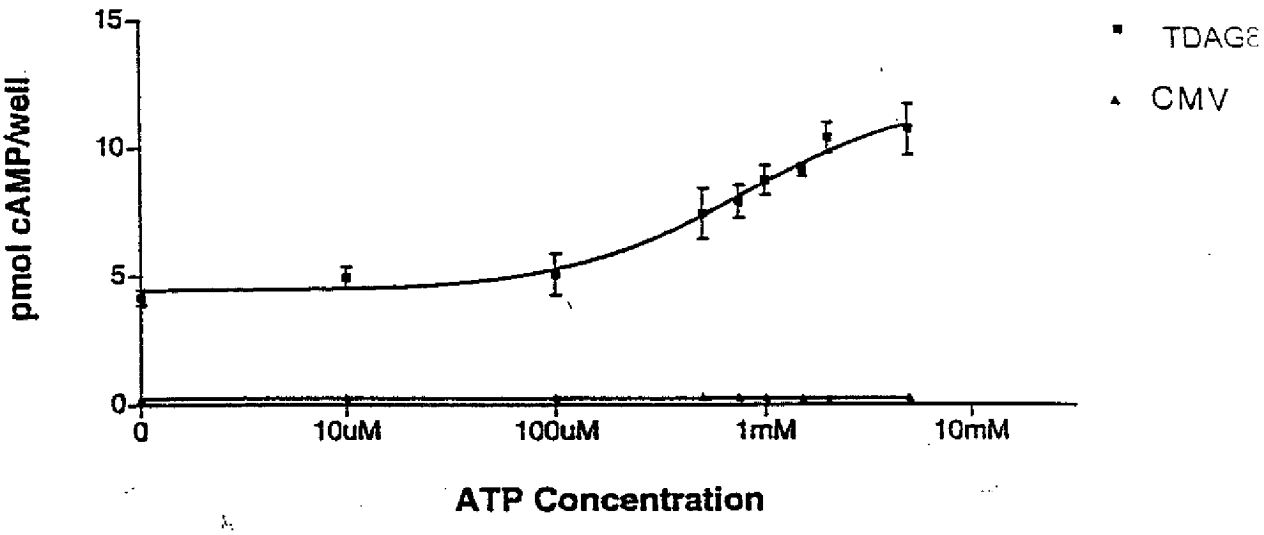


FIGURE 4A

293 Cell-based cAMP Assay

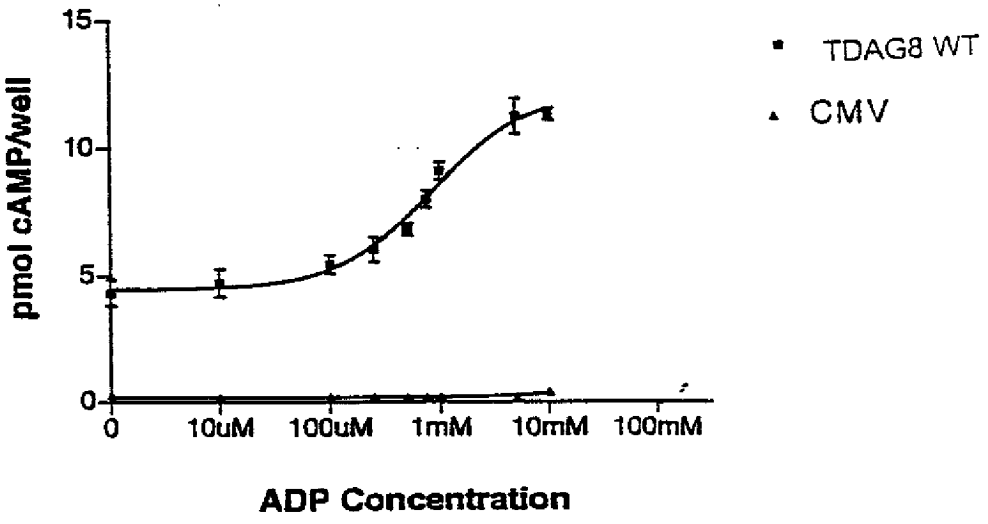


FIGURE 4B

FIGURE 5A

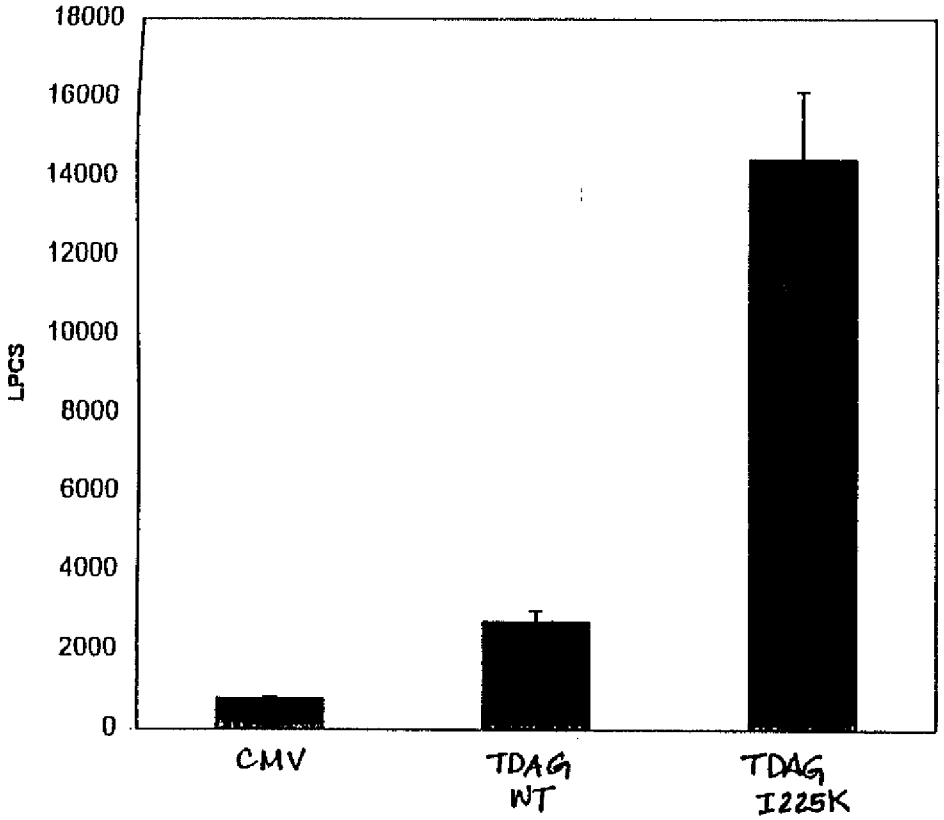
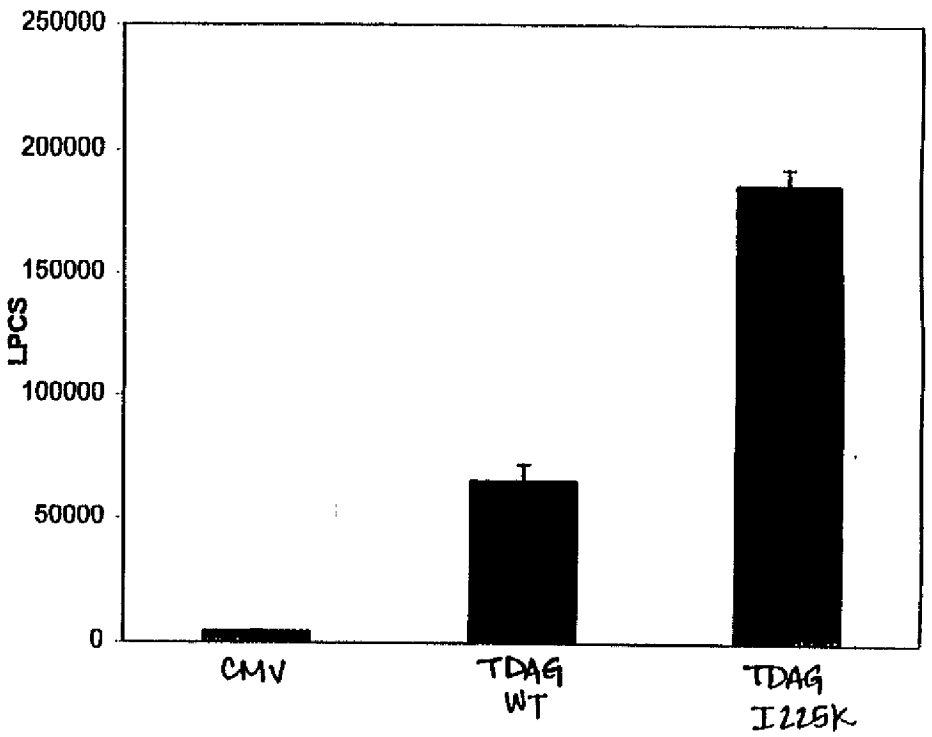


FIGURE 5B



8 x CRE-luc Reporter

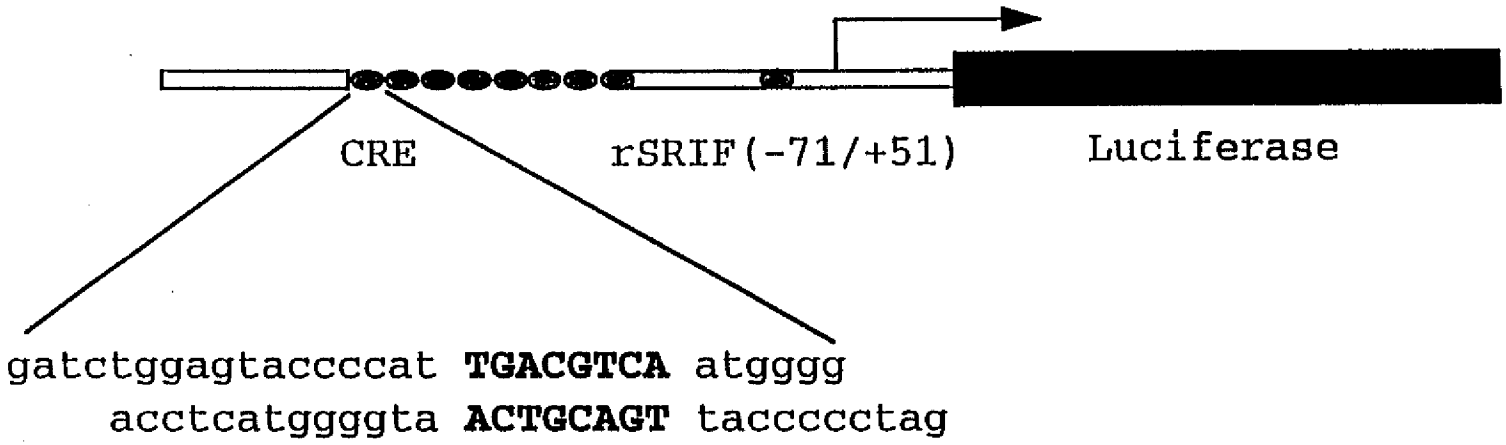
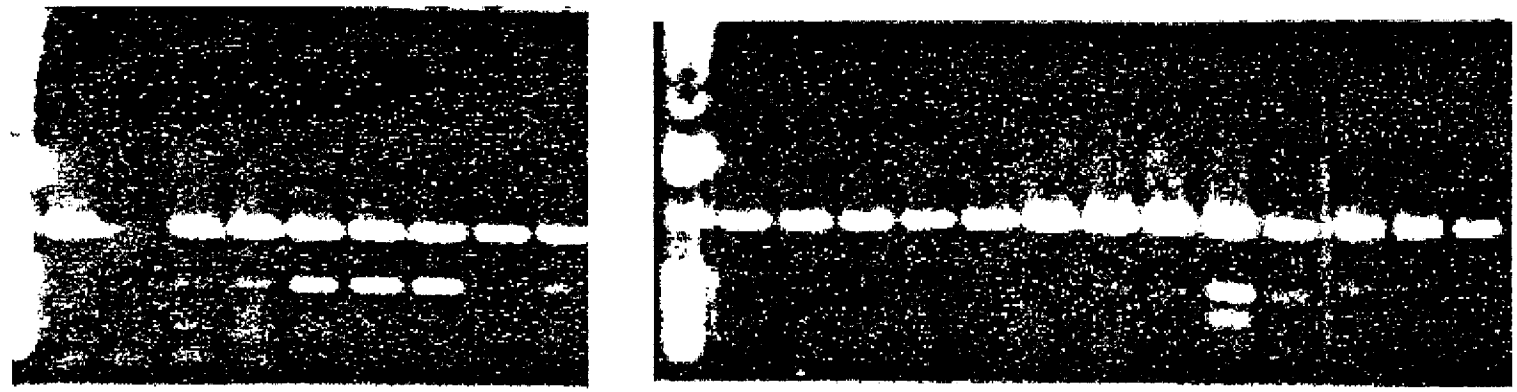


FIGURE 6

Figure 7

TDAG8 Tissue Distribution



+ pos. control
- neg. control
1) bone marrow [+]
2) fetal liver [+]
3) lymph node [++]
4) peripheral blood leukocytes [++]
5) spleen [++]
6) thymus [+]
7) tonsil [+]
8) colon [-]
9) ovary [+]
10) prostate [-]
11) small intestine [-]
12) testis [-]
13) brain [+]
14) heart [+]
15) kidney [+]
16) liver [++]
17) lung [+]
18) pancreas [+]
19) placenta [+]
20) skeletal muscle [-]